

Explainable AI-Based Hybrid LightGBM–LSTM for Banking Fraud Detection

Rafif Ramadhan Al Yarda^{1)*}, Fauziah²⁾

¹⁾²⁾Master of Information Technology, Universitas Nasional, Jakarta, Indonesia

¹⁾rafif.alyarda99@gmail.com, ²⁾fauziah@civitas.unas.ac.id

Submitted : Aug 5, 2026 | Accepted : Sep 20, 2026 | Published : Oct 1, 2026

Abstract: Banking fraud detection faces changing transaction behaviors, class imbalance, and the need to interpret the outputs of complex models. This study develops a Fraud Detection and Analysis System (FDAS) as a *Decision Support System* in the form of a *prototype* that integrates tabular and temporal modeling, imbalance handling, operational validation, and post-prediction interpretation. A quantitative experimental approach was employed using a synthetic transaction dataset constructed based on the characteristics of observed fraud and normal patterns, comprising 50 accounts, consisting of 25 fraud accounts and 25 normal accounts, covering the period from January to December 2025. *Account-level splitting* was used to prevent historical information leakage across accounts, while TPH-SMOTE and KHOI-SMOTE were applied only to the *training set*. LightGBM and LSTM were trained independently to represent tabular and temporal data, respectively, and their probabilities were subsequently combined through weighted decision-level late-fusion, with SOP serving as the *operational validation layer*. SHAP was employed as a *post-hoc explanation* to support *interpretability*, *traceability*, and *human oversight*. Across 1,000 paired evaluation units, the final configuration achieved an Accuracy of 97.4%, Precision of 96.8%, Recall of 96.2%, F1-score of 96.5%, and ROC-AUC of 0.984. Comparison across configurations indicated that the highest performance was achieved by the integrated configuration within the experimental environment used. FDAS demonstrates end-to-end integration to support fraud investigation through recommendations and interpretive information, while the final decision remains with the analyst. The use of synthetic data, 50 accounts, and a *prototype* status limits direct generalization to operational banking populations and production readiness.

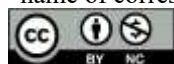
Keywords: Fraud Detection; LightGBM; LSTM; Ensemble Late-Fusion; Explainable Artificial Intelligence

INTRODUCTION

Digital transformation in the banking sector has increased the volume, diversity, and speed of transactions, making fraud patterns increasingly dynamic and not always identifiable through a single indicator. *Rule-based detection* remains relevant for known fraud patterns, but its adaptability is limited when transaction characteristics change and relationships among variables become increasingly complex. The development of *machine learning* and *deep learning* provides data-driven approaches for learning transaction patterns, including tabular characteristics and temporal dependencies that are difficult to represent through static rules (Chen et al., 2025; El Hachimi et al., 2026; Uddin, 2025). This development has made *data-driven* approaches increasingly relevant for supporting modern fraud detection, although they do not replace all functions of rule-based mechanisms.

Another challenge arises from class imbalance, particularly because fraud distribution at the transaction or *sequence* level may differ from the distribution at the account level. This condition can affect the model's ability to learn the minority class and requires *resampling* and evaluation protocols that preserve the separation between training and evaluation data (Bernardo & Della Valle, 2022; Du et al., 2026; Liu et al., 2026; Widodo et al., 2024). At the same time, predictive capability alone does not fully meet the requirements of a *fraud detection* system in an investigative context. Complex ML/DL models require outputs that can be understood and traced, while operational decision-making requires validation support and *human oversight*. XAI, particularly SHAP, has become increasingly relevant for supporting *interpretability*, *transparency*, and *auditability* in AI-based decision systems (Alshudukhi et al., 2025; Nimmy et al., 2025; Zhou et al., 2023).

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Previous studies have demonstrated substantial progress in predictive modeling through tabular models, temporal models, and *ensemble* approaches. However, the integration of tabular and temporal modeling, *imbalanced data* handling, operational validation, and *post-hoc explainability* within a single *decision support* framework still requires further development. This gap does not indicate the absence of previous approaches, but rather reflects the need to connect several components that have often been investigated separately.

Based on this gap, this study develops the **Fraud Detection and Analysis System (FDAS)** as a *Decision Support System* that integrates tabular and temporal modeling, *imbalanced data* handling, *operational SOP validation*, and SHAP as a *post-hoc explanation* mechanism. FDAS is developed as an *operational prototype* within a controlled experimental environment based on synthetic data constructed from the characteristics of observed fraud and normal transactions, rather than as a direct representation of the population of banking transactions. This study aims to develop a decision-support framework that connects predictive capability, operational validation, and *interpretability* to support fraud detection and investigation without replacing the final decision of the analyst.

LITERATURE REVIEW

Fraud Detection System

Fraud detection systems initially relied heavily on *rule-based* approaches to identify known transaction patterns. This approach remains relevant as an operational mechanism, but its flexibility is limited when fraud patterns change and relationships among variables become increasingly complex. The development of *machine learning* enables transaction patterns to be learned directly from data, while *deep learning* extends this capability through more complex representations, including temporal characteristics in transaction sequences. Studies on *deep learning* in *financial fraud detection* have demonstrated substantial progress in data modeling while also highlighting challenges related to model complexity, class imbalance, and model deployment in real-world environments (Chen et al., 2025). The use of LSTM in fraud detection also demonstrates the relevance of sequential modeling for capturing changes in transaction behavior over time (El Hachimi et al., 2026).

These developments have driven FDS toward a more *data-driven* approach, but predictive capability alone is not always sufficient for investigative requirements. Modern systems also need to provide outputs that can be traced, understood, and used to support decision-making. The role of AI in *financial crime* prevention places predictive analytics within an operational context involving validation, oversight, and human decision-making (Uddin, 2025). In this context, *explainability* helps users understand the basis of predictions and supports the traceability of analytical results in *financial fraud detection* (Zhou et al., 2023).

Previous studies have demonstrated substantial advances in *classification performance* through *machine learning* and *deep learning*, but the integration of predictive modeling, operational validation, and *post-hoc explainability* within a single *decision support* framework still requires further development (Chen et al., 2025; El Hachimi et al., 2026; Uddin, 2025). Based on this gap, this study develops the **Fraud Detection and Analysis System (FDAS)** as a *Decision Support System* that combines *predictive modelling*, *operational validation*, and *post-hoc explainability* to support investigative recommendations without replacing the analyst's role as the final decision-maker.

Hybrid LightGBM–LSTM

LightGBM and LSTM have complementary characteristics for transaction analysis. LightGBM is suitable for modeling nonlinear relationships in tabular and structured data with good modeling efficiency (Kim & Kim, 2025; Zhou et al., 2025), whereas LSTM is designed to capture temporal dependencies in sequential data and has been used in fraud detection to represent changes in transaction patterns and address class imbalance (El Hachimi et al., 2026). These differences in representation enable transaction attributes and behavioral dynamics to be analyzed through appropriate modeling pathways, thereby reducing the risk of mismatch between input structure and model characteristics.

The integration of the two models can be performed at the feature, parameter, or decision level. This study employs **decision-level late-fusion**, in which LightGBM and LSTM are trained independently on their respective representations, and their prediction probabilities are subsequently combined after inference. This approach preserves the characteristics of each model and provides flexibility when heterogeneous models operate on different data structures. The use of ensembles combining sequential models and LightGBM has also been demonstrated in previous research (Gong et al., 2025). Accordingly, *late-fusion* is selected not because it is always superior to a monolithic architecture, but because it is consistent with the research design, which maintains two independent modeling pathways.

In the context of fraud detection, the combination of LightGBM and LSTM enables structured transaction characteristics and temporal behavioral changes to be analyzed complementarily. The literature demonstrates the development of LightGBM, LSTM, and ensemble approaches for predictive modeling, while the integration of heterogeneous model outputs into a *decision support* framework that also supports operational validation and *post-*

hoc explainability still requires further development (El Hachimi et al., 2026; Gong et al., 2025; Kim & Kim, 2025). This study positions *late-fusion* as the integration mechanism for two heterogeneous models without forming a single joint learning architecture.

Explainable Artificial Intelligence (XAI)

The increasing complexity of *machine learning* and *deep learning* models makes predictive outputs increasingly difficult to trace based solely on classification results. In the context of *financial fraud detection*, interpretability is relevant because risk predictions need to be understood by analysts so that they can be reviewed and used as a basis for investigation. XAI therefore not only concerns providing explanations of models but also supports transparency in the analytical process and traceability of decisions in systems with operational consequences (Alshudukhi et al., 2025; Zhou et al., 2023).

One approach used for this purpose is *SHapley Additive exPlanations* (SHAP), a *post-hoc* method that explains the relative contribution of features to model outputs without being part of the learning process. SHAP provides a **local explanation** for understanding feature contributions to an individual prediction and a **global explanation** for summarizing feature contributions across a set of predictions (Nimmy et al., 2025; Zhou et al., 2023). In the context of the models used in this study, **TreeSHAP** is relevant for tree-based models such as LightGBM, whereas **DeepSHAP** can be used to explain *deep learning* models such as LSTM. Both approaches are used to understand model outputs, rather than to alter the classification process or improve predictive metrics.

From a decision support system perspective, *explainability* can strengthen the relationship **prediction** → **explanation** → **analyst interpretation** → **audit/investigation**. This requirement is increasingly relevant in AI systems that require transparency, traceability, and *human oversight* (Alshudukhi et al., 2025; Nimmy et al., 2025). The literature has demonstrated the benefits of XAI in helping users understand model decisions, but the integration of *explainability* as part of a *decision support* workflow in *fraud detection* that utilizes tabular and temporal models still requires further development (Zhou et al., 2023). Based on this gap, this study positions SHAP as a *post-hoc explanation layer* in FDAS to support *interpretability* and *traceability* without altering model probabilities or the *FinalScore*.

Imbalanced Datasets in Fraud Detection

Class imbalance is an important issue in *fraud detection* because the distribution of observations may be dominated by normal transactions rather than fraud. This condition can make models more likely to learn majority-class patterns and reduce sensitivity to the minority class. Imbalance needs to be assessed at the level of observations used for learning, rather than solely based on the number of accounts, because a single account may generate different numbers of transactions or *sequences*. Therefore, balance at the account level does not automatically indicate balance at the transaction or *sequence* level. Handling this issue is generally performed through *resampling*, such as **Random Oversampling (ROS)**, which adds duplicated minority observations; **SMOTE**, which generates synthetic samples based on relationships among minority samples; and **SMOTE-ENN**, which combines *oversampling* with observation cleaning to help improve class separation. ROS is relatively simple but can increase duplication and the risk of *overfitting*, whereas SMOTE and its variants are more adaptive but remain sensitive to data structure and *noise* (Bernardo & Della Valle, 2022; Widodo et al., 2024).

The development of *oversampling* methods has led to approaches that are more adaptive to data characteristics. TPH-SMOTE combines the SMOTE process with heuristic and optimization mechanisms to generate a more targeted representation of the minority class (Du et al., 2026), whereas KHOI-SMOTE utilizes cluster structure and *h-outlyingness* in the generation of synthetic samples (Liu et al., 2026). These approaches do not guarantee improvements across all metrics because their effects depend on data distribution and model characteristics. Resampling can also increase sensitivity to the minority class, but changes in distribution may affect Precision, Recall, and F1-score. The risk of *data leakage* arises when *oversampling* is performed before *data splitting*, because synthetic samples may be related to their original observations, which are subsequently distributed across evaluation subsets and may produce overly optimistic performance estimates. Therefore, *resampling* should be applied only to the training data, while *validation* and *test* sets should retain their original distributions (Widodo et al., 2024).

In this study, TPH-SMOTE and KHOI-SMOTE are applied to the **training set** after data splitting, while the *validation* and *test* sets are maintained without *resampling*. This protocol preserves the separation between learning and evaluation and reduces the risk of *data leakage*. Because transaction distributions may differ from account-level distributions, the selection of the *resampling* methods is adjusted to the observation unit and model characteristics used in *fraud detection*.

METHOD

Research Design

This study employs a quantitative experimental approach to develop the **Fraud Detection and Analysis System (FDAS)** as a *Decision Support System* to support the detection and investigation of transaction fraud in the banking sector. The study uses a synthetic dataset constructed based on the characteristics of observed fraud and normal transaction patterns from several conventional banks in Indonesia. The dataset consists of **50 accounts**, comprising 25 fraud accounts and 25 normal accounts, with transaction histories from January to December 2025. The dataset is used as a controlled experimental environment and is not intended to represent actual customer transactions or the entire banking population.

Following *preprocessing*, the data are split at the account level (*account-level splitting*) to ensure that the entire transaction history of a single account remains within the same subset and does not result in *account-history leakage* across subsets. The separated data are then used in two independently trained modeling pathways: **LightGBM** for tabular transaction representation and **LSTM** for temporal transaction sequences (Yousefimehr & Ghatee, 2025). In this study, the term *hybrid* refers to the integration of heterogeneous models through **decision-level late-fusion**, without *joint learning* or *parameter sharing*. The output probabilities of the two models are then **integrated with the SOP validation score through weighted decision-level late-fusion** with weights of 0.40 for LightGBM, 0.40 for LSTM, and 0.20 for SOP. SOP is positioned as an *operational validation layer* after the model outputs become available and is used in the formation of the *FinalScore*, while the weighting configuration represents the conceptual design of the study rather than the result of empirical optimization.

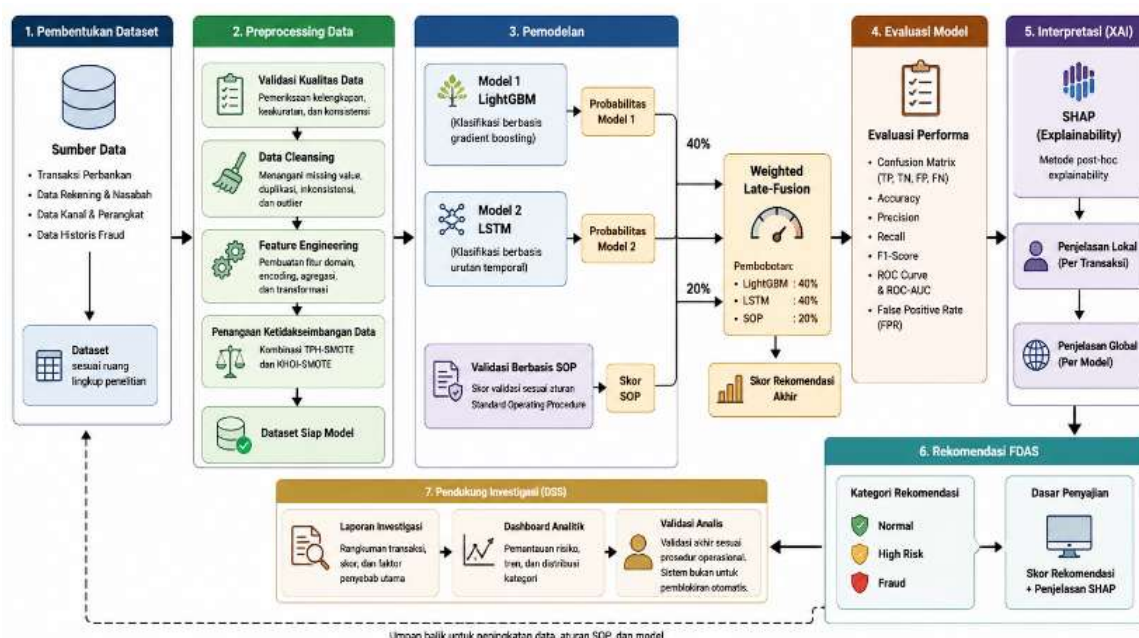


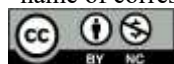
Figure 1. Fraud Detection and Analysis System (FDAS) Workflow

The workflow clearly separates the decision path from the interpretation path. In the *decision path*, the outputs of LightGBM and LSTM are combined with SOP validation to produce the *FinalScore*, which is subsequently passed to the *recommendation layer*. In the *explanation path*, model predictions are analyzed using **SHAP** in a *post-hoc* manner to generate feature contribution information without altering the model probabilities, SOP score, or *FinalScore* (Zhou et al., 2023). Thus, **SOP serves to validate the decision from an operational perspective, whereas SHAP explains the basis of the prediction for the analyst**. FDAS is positioned as a *prototype* to demonstrate this *end-to-end* integration; system recommendations are used as **supporting information for investigation**, while the final decision remains with the analyst or human decision-maker.

Dataset and Research Variables

This study uses a synthetic transaction dataset as a controlled experimental environment for the development and testing of FDAS. The dataset was constructed through scenario-based synthetic generation based on the characteristics and patterns of fraud and normal transactions obtained through observations at several conventional banks in Indonesia, without using or publishing actual customer transactions. The generation scenarios preserve variations in transaction values, time, activity frequency, transaction destination patterns, remarks, and balance

*name of corresponding author



changes so that normal and fraudulent behavioral characteristics can be conceptually represented. Class scenarios were predetermined at the account level, after which transaction attributes were generated according to the corresponding behavioral profiles, while the target label remained separated from the predictor variables. The dataset consists of 50 accounts, comprising 25 fraud accounts and 25 normal accounts, with transaction histories from January to December 2025. This configuration was used as a controlled experimental environment and was not intended to directly represent the population of banking transactions (Almusallam & Qayyum, 2025).

Table 1. Dataset Composition and Unit of Analysis

Category	Description
Dataset type	Synthetic transaction dataset
Domain	Banking transaction fraud detection
Number of accounts	50 accounts
Fraud accounts	25 accounts
Normal accounts	25 accounts
Transaction period	January–December 2025
Account-level unit	Account
Transaction-level unit	Individual transaction
Sequential unit	Chronologically ordered transaction sequence
Target label	Fraud / Normal
Real customer transaction data	Not used

At the analytical level, the account serves as the unit for behavioral group formation and data splitting, the transaction serves as the tabular observation unit for LightGBM, while the sequence serves as the temporal observation unit for LSTM. A single account may contain multiple transactions and sequences; therefore, 50 accounts do not correspond to 50 transaction observations, 50 sequences, or 50 evaluation units. The fraud/normal label is defined as the target variable, while predictor attributes are constructed from transaction and behavioral characteristics that are separated from the label. The label is not included as a feature, and the rules used to construct the label are not deterministically replicated in the predictors; this procedure is designed to reduce the risk of direct target leakage.

The generation of synthetic observations follows behavioral scenarios that distinguish normal and fraudulent characteristics without specifying undocumented probability distributions or empirical thresholds. A summary of the construction scenarios is presented in Table 2.

Table 2. Synthetic Transaction Generation Scenario

Data Aspect	Normal Scenario	Fraud Scenario
Transaction Value	Relatively consistent with account activity	Irregular or abrupt value variation
Timestamp	Distributed within expected activity periods	Unusual temporal concentration or deviation
Frequency	Consistent transaction intensity	Elevated or abnormal transaction bursts
Destination Pattern	Recurrent or expected destinations	Unusual or changing destinations
Remark	Consistent descriptive patterns	Atypical or inconsistent descriptions
Balance	Changes compatible with transaction activity	Abrupt or atypical balance changes

The table represents a synthetic data construction scenario, not an empirical distribution or a set of FDAS classification rules. The research variables include numerical, temporal, textual, and categorical attributes. Debit and Credit represent transaction values, Timestamp represents temporal characteristics, Remark represents transaction descriptions, Balance represents changes in account balance, Frequency represents activity intensity, and Destination Pattern represents transaction destination behavior (Yousefimehr & Ghatee, 2025).

Table 3. Research Variables

Variable	Data Type	Analytical Function
Debit	Numerical	Represents outgoing transaction value
Credit	Numerical	Represents incoming transaction value
Timestamp	Date and time	Represents temporal characteristics of transactions
Remark	Text	Represents transaction description

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Balance	Numerical	Represents account balance change
Frequency	Numerical	Represents transaction activity intensity
Destination Pattern	Categorical	Represents transaction destination behavior

These variables are used to form tabular representations for LightGBM and sequential representations for LSTM. The use of synthetic data limits the external validity of the study but enables methodological evaluation within a controlled environment without exposing sensitive customer information. Therefore, the dataset is positioned as the basis for testing the integration of the FDAS methodology and is not intended to comprehensively represent operational banking conditions.

Data Preprocessing

Preprocessing is performed to produce a consistent analytical dataset before modeling. The initial stage includes checking for missing values, duplicates, invalid data, and consistency of timestamps, transaction values, and remarks. Operations that do not require learning parameters are performed as part of the cleaning process, whereas transformations that depend on data-derived parameters are determined from the training set and applied to the validation and test sets without refitting. This procedure is designed to limit the risk of preprocessing leakage before model learning is performed.

The dataset is then split based on Account ID with a target proportion of 70% for training, 15% for validation, and 15% for testing. All transaction histories from a single account are placed in only one subset so that the history of the same account is not distributed across the training, validation, and test sets. Because the study uses 50 accounts, the ratio is treated as a target allocation at the account level, while the actual number of accounts follows integer-based allocation. Randomization is applied during account allocation, but the initial random seed was not documented and therefore was not retrospectively reconstructed. The stratification configuration was also not reconstructed if it was not recorded in the initial experimental configuration.

Table 4. Data Splitting and Preprocessing Protocol

Stage	Unit/Scope	Procedure	Leakage Control
Cleaning	Synthetic transactions	Validation and cleaning	Invalid/duplicate records handled
Account-level splitting	Account ID	Target 70:15:15	Account histories not shared across subsets
Training preprocessing	Training set	Parameters learned from training	No fitting on validation/test
Validation/Test transformation	Validation/Test	Apply training-derived transformation	No refitting
Oversampling	Training only	TPH-SMOTE + KHOI-SMOTE	No oversampling on validation/test
Temporal sequencing	Account history	Chronological ordering	No cross-account sequence mixing
Model input	Processed data	Tabular → LightGBM; Sequence → LSTM	Separate model representations

After splitting, TPH-SMOTE and KHOI-SMOTE are applied only to the training observations (Du et al., 2026; Liu et al., 2026). Validation and test sets do not undergo oversampling so that their evaluation distributions remain unmodified by synthetic resampling. This approach preserves the separation between learning and evaluation data and limits the risk of resampling leakage. Transactions within each account are subsequently ordered chronologically to form LSTM sequences, while tabular representations are used by LightGBM (El Hachimi et al., 2026; Yousefimehr & Ghatee, 2025). Boundaries between accounts are maintained during sequence formation so that histories from different accounts are not mixed.

Within the analytical structure of the study, the account functions as the unit for history grouping and data splitting, the transaction functions as the tabular observation unit, the sequence functions as the temporal observation unit, while the evaluation unit follows the model outputs used in metric calculations. Therefore, the number of accounts is not equated with the number of transactions, sequences, or evaluation units. The number of transactions and sequences in each subset follows the structure of the constructed dataset and is not reconstructed as a new figure in this subsection.

Hybrid LightGBM–LSTM–XAI Architecture

FDAS uses **LightGBM** to process tabular transaction representations and **LSTM** to model transaction sequences chronologically. The two models are trained independently without *joint learning*, *parameter sharing*, or parameter exchange, and subsequently generate fraud probabilities separately. In this study, the term *hybrid* refers to the integration of heterogeneous models through **decision-level late-fusion**, allowing each model to retain an input representation appropriate to its data characteristics (Yousefimehr & Ghatee, 2025). The main experimental configuration is summarized in Table 5.

Table 5. Main Configuration of the LightGBM and LSTM Models

Parameter	LightGBM	LSTM
Input representation	Tabular transaction features	Chronological transaction sequences
Learning rate	0.05	0.001
Core architecture	200 estimators; 31 leaves	2 LSTM layers; 64 hidden units
Dropout	Not applicable	0.20
Batch size	Model-dependent	32
Epochs	Model-dependent	50
Optimizer	Boosting-based optimization	Adam
Output	Fraud probability	Fraud probability
Training setup	Independent training	Independent training

This configuration is used as a controlled experimental setting and does not represent the result of empirical optimization. The output probabilities of the two models are then integrated with the operational validation score through *weighted decision-level late-fusion*:

$$P_{final} = (0.4 \times P_{LightGBM}) + (0.4 \times P_{LSTM}) + (0.2 \times S_{SOP}) \quad (1)$$

Here, $P_{LightGBM}$ and P_{LSTM} represent the respective fraud probabilities, while S_{SOP} represents the operational validation score. The 40:40:20 weighting is a **conceptual design** and is not positioned as an optimal weighting. SOP is applied after the model probabilities become available and is not used as a *predictor*, training feature, target, or third classifier. The five validation indicators used are summarized in Table 6.

Table 6. SOP Validation Indicators and Rule Configuration

Indicator	Weight	Training Use
Transaction Velocity Check	0.25	No
Geographic Anomaly Check	0.15	No
Transaction Frequency Anomaly	0.20	No
Destination Pattern Check	0.20	No
Transaction Amount Deviation	0.20	No

Each indicator produces a normalized operational score within the range of 0–1, which is aggregated into S_{SOP} according to the defined rule configuration. These indicators originate from operational and investigative examination requirements within the scope of the study, rather than from the fraud/normal labels. This functional separation is designed to prevent *SOP-derived* information from entering predictor construction, target formation, and model training. The *FinalScore* is subsequently mapped to decision categories based on the thresholds defined in Table 7.

Table 7. FDAS Decision Thresholds Based on FinalScore Values

FinalScore Range	Risk Category	System Recommendation
0.00–<0.40	Normal	Activity is considered normal; no further investigation is required
0.40–<0.70	High Risk	Further monitoring and validation by the analyst
0.70–1.00	Fraud	Investigation is recommended in accordance with operational procedures

SHAP is positioned within the **post-prediction interpretation path** and is not part of *FinalScore* formation. The decision path follows **LightGBM probability + LSTM probability + SOP validation → weighted late-fusion → FinalScore → recommendation**, whereas the interpretation path follows **Model Prediction → SHAP → Feature Contribution → Analyst Interpretation**. Thus, SHAP does not alter the model probabilities, S_{SOP} or

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

FinalScore, and is not used during training (Zhou et al., 2023). The architecture is summarized in Figure 2, with the final decision remaining with the analyst.

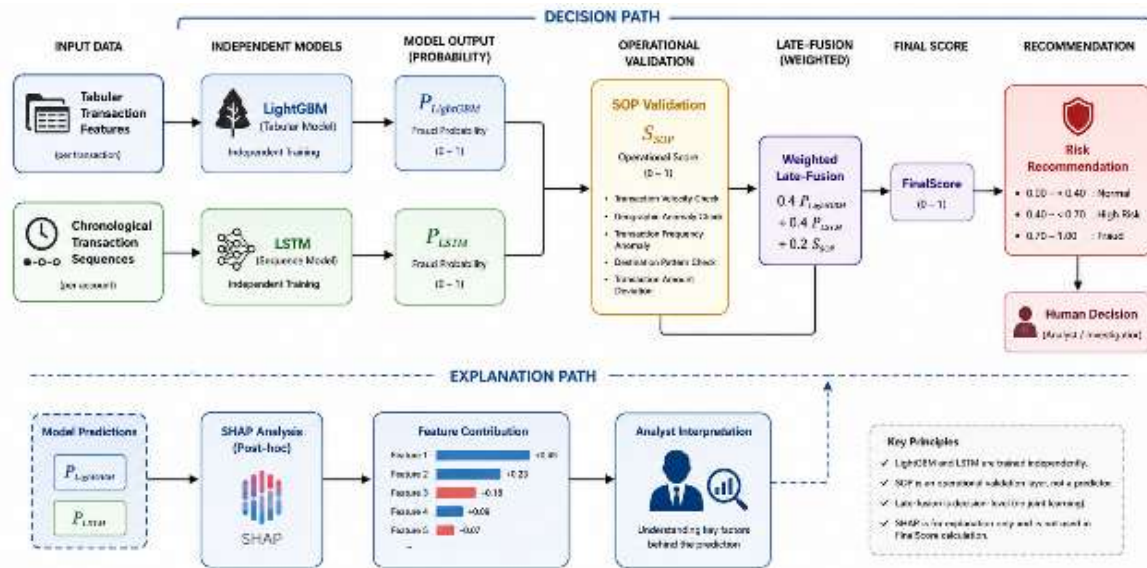


Figure 2. Ensemble Late-Fusion Architecture of the Proposed FDAS

Explainable Artificial Intelligence (XAI)

This study uses **Explainable Artificial Intelligence (XAI)** based on *SHapley Additive exPlanations* (SHAP) as a post-prediction interpretation mechanism in FDAS. SHAP is applied after inference to explain the fraud probabilities generated by each model. **TreeSHAP through TreeExplainer** is used for LightGBM, whereas **DeepSHAP through DeepExplainer** is used for LSTM. The *background/reference* is derived exclusively from the *training partition*, with a maximum of 50 tabular observations for LightGBM and 50 *sequences* for LSTM; if fewer training units are available than these limits, all available units are used. This configuration represents a **reconstructed methodological configuration** for the controlled experimental environment and is not the result of empirical optimization. Validation and test data are not used as *background/reference* data.

Table 8. XAI Configuration

Component	LightGBM	LSTM
Explainer	TreeSHAP / TreeExplainer	DeepSHAP / DeepExplainer
Input	Tabular observation	Chronological transaction sequence
Background	Up to 50 training observations	Up to 50 training sequences
Local explanation	Feature contribution	Timestep-feature contribution
Global explanation	Mean Absolute SHAP	Aggregated Mean Absolute SHAP

Local explanation is used to explain feature contributions in a single observation or sequence, whereas global explanation summarizes the relative contributions of features across the analyzed observations. In LSTM, SHAP values at the timestep–feature level are aggregated into feature-level contributions by summarizing the absolute magnitude of contributions across all timesteps. Global contributions are subsequently summarized using **Mean Absolute SHAP**; for LightGBM, the calculation is performed directly from the SHAP values of tabular observations, whereas for LSTM, it uses feature contributions obtained through temporal aggregation. The global contribution rankings are reported separately for each model to preserve attribution specific to the data representation used. Higher values indicate greater relative contributions to the model output and are not interpreted as causal relationships (Zhou et al., 2023).

*name of corresponding author



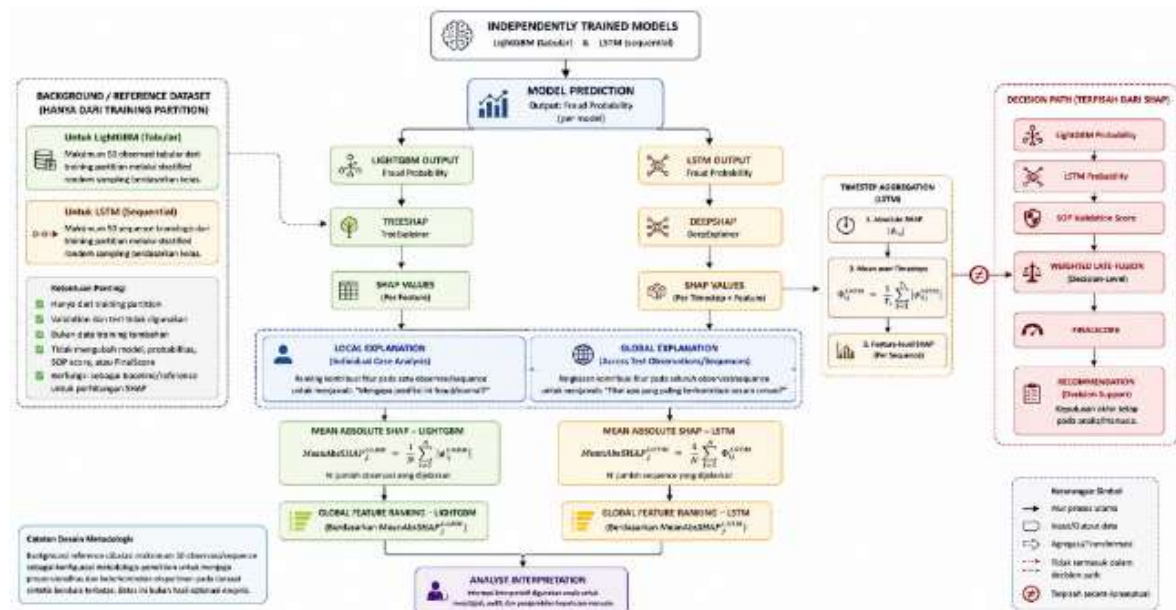


Figure 3. Post-hoc SHAP Explanation Workflow in FDAS

SHAP is positioned within the interpretation path and is not part of the decision path. The **decision path** follows *LightGBM Probability + LSTM Probability + SOP Score* → *Weighted Late-Fusion* → *FinalScore* → *Recommendation*, whereas the **explanation path** follows *Model Prediction* → *SHAP* → *Feature Contribution* → *Analyst Interpretation*. Thus, SHAP is not used during *training*, does not serve as a *predictor* or *classifier*, and does not alter the model probabilities, S_{SOP} , or *FinalScore*. This analysis is positioned as **SHAP explanation analysis**, rather than *SHAP validation*, because the study does not conduct formal testing of explanation stability or quality. This layer supports *interpretability*, *transparency*, *traceability*, and *auditability* while maintaining *human oversight* (Lukács & Váradi, 2023; Zhou et al., 2023).

Model Evaluation

Evaluation is conducted after *account-level splitting*, with the **account as the splitting unit**, **transaction as the LightGBM observation unit**, and **sequence as the LSTM observation unit**. All transaction histories belonging to a given account are placed in a single subset so that historical overlap across subsets can be controlled. In the *weighted decision-level late-fusion* configuration, the LightGBM and LSTM outputs that can be paired with the SOP score are aligned into **paired decision units** for final evaluation. This configuration produces **1,000 paired evaluation units** used to construct the *confusion matrix*. This number does not represent 1,000 accounts; the 50 accounts remain the research and splitting units, while a single account may generate multiple *transactions* and/or *sequences* that become evaluation units after the splitting process.

Classification performance is determined from the *confusion matrix*, namely **True Positive (TP)** for fraud correctly detected, **True Negative (TN)** for normal observations correctly classified, **False Positive (FP)** for normal observations incorrectly flagged as fraud, and **False Negative (FN)** for fraud that is not detected. Thus, in the final evaluation configuration, the number of evaluation units satisfies **. The metrics used include Accuracy, Precision, Recall, F1-score, ROC-AUC, AUC-PR, and FPR. Accuracy indicates overall correctness, Precision indicates the proportion of fraud predictions that are correct, Recall indicates the proportion of actual fraud cases that are detected, F1-score summarizes the balance between Precision and Recall, ROC-AUC assesses discriminative ability across different *thresholds*, AUC-PR summarizes the Precision–Recall relationship for the fraud class, while FPR measures the proportion of normal observations incorrectly flagged as fraud. Recall is prioritized because **False Negative represents undetected fraud**, while the other metrics are used as complementary measures.

The primary metrics are calculated as follows:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (4)$$

$$F1 = 2 \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

*name of corresponding author



For the final evaluation, **TP=358**, **TN=616**, **FP=12**, and **FN=14**, resulting in:

$$358 + 616 + 12 + 14 = 1.000$$

Dan:

$$Accuracy = \frac{358 + 616}{1.000} = 0.974 = 97.4\%$$

Thus, the *confusion matrix* and the Accuracy value of 97.4% can be mathematically traced from the 1,000 *paired evaluation units*. Validation and test data do not undergo *oversampling*; **TPH-SMOTE** and **KHOI-SMOTE** are applied only to the *training set*. Therefore, the metrics are calculated on evaluation data that have not been modified by *synthetic resampling*. The evaluations of LightGBM, LSTM, and *weighted decision-level late-fusion* are maintained as three distinct levels of analysis according to their respective evaluation units.

RESULT

Data Cleansing and Preprocessing Results

The *data cleansing* and *preprocessing* stages were applied to the *transaction records* before model input formation. The initial dataset consisted of **15,240 transaction records** from 50 accounts over the period from January to December 2025. The cleaning process was performed at the transaction level, so the number of accounts remained 50, with a composition of 25 fraud accounts and 25 normal accounts. This 25:25 composition applies at the account level and does not indicate that the numbers of fraud and normal transactions were distributed in equal proportions.

A total of **89 transaction records** were affected by *missing values*, comprising **114 missing cells**, with no records removed at this stage. Subsequently, **83 duplicate records** were removed, reducing the dataset to **15,157 records**. A total of **7 records** had format inconsistencies and were corrected without reducing the number of observations. Outlier examination then identified **231 records as outlier candidates**. After validity assessment, **44 candidates** were removed because they were categorized as invalid extreme observations or *noise*, while **187 candidates** were retained because they remained valid based on their transaction characteristics. Thus, the 231 candidates were not equated with the number of removals; only 44 records were actually removed during the *outlier treatment* stage.

Table 9. Summary of Data Cleansing Results

Cleansing Stage	Affected Records	Missing Cells / Candidates	Records Removed	Records Remaining
Raw transaction dataset	-	-	-	15.240
Missing-value handling	89	114	0	15.240
Duplicate detection	83	-	83	15.157
Invalid-format correction	7	-	0	15.157
Outlier assessment	231	231 candidates	44	15.113
Final cleaned dataset	-	-	-	15.113

After all *cleansing* stages were completed, the dataset consisted of **15,113 transaction records**, or approximately **99.17%** of the initial data. The process was intended to maintain the statistical and structural consistency of the synthetic dataset with the transaction characteristics underlying its construction. The cleaned dataset was subsequently used to construct features and input representations for LightGBM and LSTM *sequences*. After *account-level splitting*, TPH-SMOTE and KHOI-SMOTE were applied only to the *training observations*, while the validation and test sets did not undergo *oversampling*. Therefore, the **15,113 cleaned transaction records** do not include synthetic observations generated during the training data balancing stage.

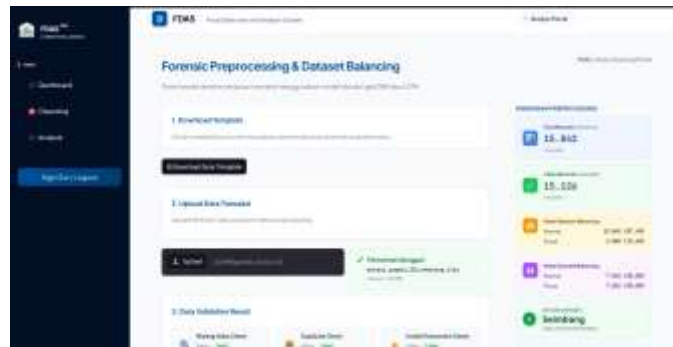


Figure 4. FDAS Cleansing Module Interface.

Figure 4 shows the implementation of the *cleansing* module in the FDAS prototype, while the quantitative cleaning results refer to Table 9. The cleaned dataset serves as the basis for constructing tabular and temporal representations before the modeling stage.

Fraud Detection and Analysis System (FDAS) Implementation Results

The implementation results position FDAS as an **operational prototype and proof-of-concept** to demonstrate the integration of *end-to-end* functions within a controlled experimental environment based on a synthetic dataset. The prototype workflow includes user authentication, data preparation, analysis, operational validation, recommendation, and reporting. The login interface and *dashboard* in **Figure 5** show the user access point and integration of the prototype's main navigation.

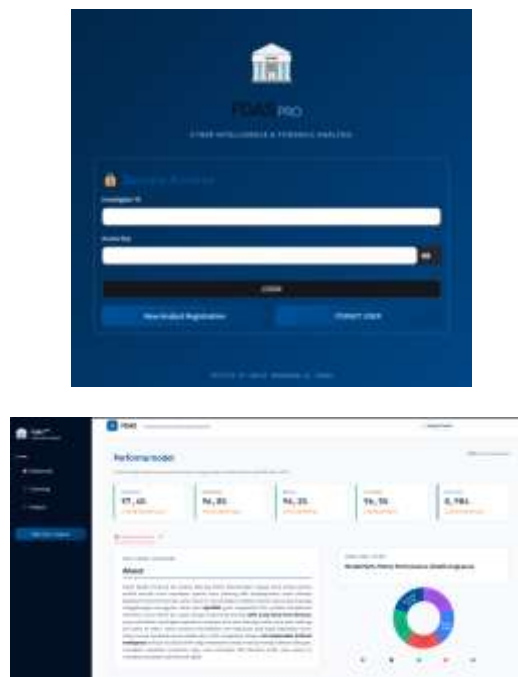


Figure 5. FDAS Login Page and Dashboard.

The interface serves as the starting point for the data processing and analysis modules. **Figure 6** illustrates the connectivity among the *cleansing*, analysis, and investigative reporting modules. Within the analytical workflow, the outputs of LightGBM and LSTM are processed through *weighted decision-level late-fusion* and SOP as the *operational validation layer*, while SHAP is applied in a *post-hoc* manner to provide feature contribution information. The integrated results are subsequently passed to the *recommendation layer* without using SHAP to modify the *FinalScore*.

*name of corresponding author



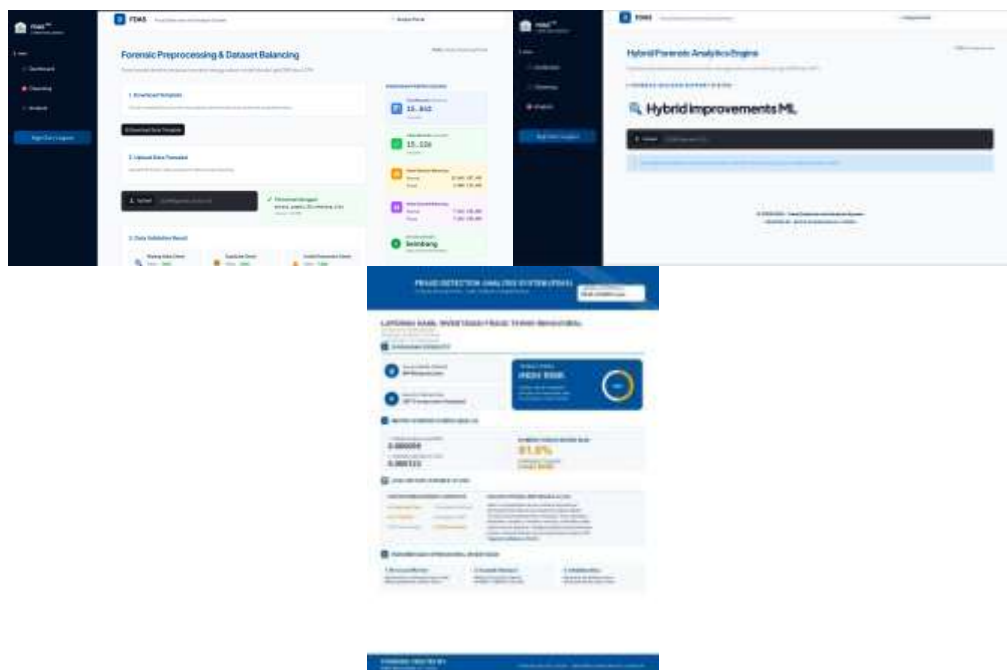


Figure 6. FDAS Cleansing, Analysis, and Investigative Reporting Prototype Interfaces

A summary of the functions demonstrated by the prototype is presented in the following table.

Table 10. FDAS Prototype Functional Demonstration

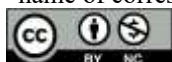
System Module	Prototype Status	Demonstrated Function
Registration/Login	Demonstrated in prototype	User authentication
Dashboard	Demonstrated in prototype	System summary and access to main modules
Data Cleansing	Demonstrated in prototype	Transaction validation and dataset preparation
Analysis	Demonstrated in prototype	Transaction analysis and model inference
Hybrid Analytics Engine	Demonstrated in prototype	LightGBM/LSTM probability integration
SOP Validation	Demonstrated in prototype	Operational validation signal
Recommendation Layer	Demonstrated in prototype	Risk category and investigation recommendation
SHAP/XAI	Demonstrated in prototype	Feature contribution explanation
PDF Investigative Report	Demonstrated in prototype	Structured analytical reporting

The prototype demonstrates the integration of the main functions, ranging from authentication and data preparation to analytical inference, operational validation, interpretation, recommendation, and reporting. **Figure 6** also shows the structured investigative report output as part of this workflow. This visual evidence demonstrates the presence and demonstration of the prototype interface functions, rather than operational validation in a real banking environment. FDAS continues to function as a *Decision Support System*; recommendations are used as supporting information for investigation, while the final decision remains with the analyst, without automatic account blocking.

Hybrid LightGBM–LSTM Model Implementation Results

The model implementation was evaluated at units appropriate to the characteristics of each model. LightGBM was evaluated at the transaction level, LSTM at the *sequence* level, while the *weighted decision-level late-fusion* configuration was evaluated on **1,000 paired evaluation units** that paired model probability outputs with operational validation. Thus, the 50 accounts represent the units for dataset formation and splitting, rather than the number of observations directly used to calculate the metrics. In the individual configurations, LightGBM achieved an *Accuracy* of 92.8%, *Precision* of 91.9%, *Recall* of 91.3%, and *F1-score* of 91.6%, whereas LSTM achieved an *Accuracy* of 94.6%, *Precision* of 94.1%, *Recall* of 93.5%, and *F1-score* of 93.8%.

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Table 11. Ablation Study of the Proposed Fraud Detection Configuration

Configuration	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	ROC-AUC	AUC-PR
LightGBM	92,8	91,9	91,3	91,6	0,950	0,918
LSTM	94,6	94,1	93,5	93,8	0,963	0,936
LightGBM + LSTM	96,1	95,4	95,1	95,2	0,975	0,954
LightGBM + SOP	94,8	94,0	93,7	93,8	0,968	0,946
LSTM + SOP	96,0	95,4	94,8	95,1	0,978	0,960
LightGBM + LSTM + SOP (Final FDAS)	97,4	96,8	96,2	96,5	0,984	0,972

Comparison across configurations shows that integrating tabular and sequential representations produces higher performance than the individual models within the compared configurations. Configurations involving SOP also show higher performance patterns across several metrics, while **Final FDAS** records the highest values across the analyzed configurations, with an *Accuracy* of 97.4%, *Precision* of 96.8%, *Recall* of 96.2%, *F1-score* of 96.5%, *ROC-AUC* of 0.984, and *AUC-PR* of 0.972. The *Accuracy* value is calculated on the model evaluation units and does not represent 50 account-level decisions. *Recall* remains prioritized as a risk-oriented metric because *False Negative* represents undetected fraud. These results are positioned as comparative patterns within a controlled experimental environment based on a synthetic dataset, rather than as evidence of universal optimality or isolated causal contributions from each component.



Figure 7. Confusion Matrix of the Final FDAS Configuration

In the final configuration, the *confusion matrix* shows 358 *True Positive*, 616 *True Negative*, 12 *False Positive*, and 14 *False Negative* cases out of 1,000 *paired evaluation units*. This distribution is consistent with an *Accuracy* of 97.4%, *Precision* of 96.8%, *Recall* of 96.2%, and *F1-score* of 96.5%. Overall, the combined configuration of LightGBM, LSTM, and SOP demonstrates the highest performance among the compared configurations, while also supporting the conceptual feasibility of *weighted late-fusion* in FDAS. These findings remain constrained by the use of a synthetic dataset and a controlled experimental environment and therefore cannot be interpreted as evidence of direct generalization to the population of banking transactions or readiness for operational *deployment*.

DISCUSSIONS

Analysis of the Hybrid LightGBM–LSTM Implementation

The difference in performance between the individual models indicates that LightGBM and LSTM utilize information characteristics that are not entirely identical. LightGBM can capture nonlinear interactions among structured transaction features, but tabular representations do not explicitly preserve sequence order and temporal dependencies. LSTM, in contrast, is more suitable for modeling behavioral changes throughout a sequence, but such representations are not always the most efficient form for capturing the heterogeneity of tabular attributes. This difference supports the concept of *complementary predictive perspectives*: the two models do not merely increase classification capacity but provide different perspectives on the same transaction patterns. This finding is consistent with the use of LightGBM on structured data (Kim & Kim, 2025) and LSTM for sequence-based fraud detection (El Hachimi et al., 2026).

Late-fusion utilizes this complementarity by combining probabilities after each model completes its inference. This mechanism preserves the representation and learning process of each model without forcing both models into a single representation space. Ensemble approaches combining temporal models and LightGBM have also demonstrated that heterogeneous models can be used complementarily (El Hachimi et al., 2026). In this study, late-fusion provides a more modular architecture than joint learning because the training, development, and control of each model can be performed separately. This consideration is relevant in a synthetic experimental environment with a limited account-level scale, as modular integration can help control the optimization complexity required

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

to learn tabular and temporal representations simultaneously. However, this consideration represents a design rationale rather than evidence that late-fusion is universally superior to joint-learning architectures.

The composition of **40% LightGBM, 40% LSTM, and 20% SOP** retains both probability sources as the primary predictive components, while SOP is positioned as a complementary operational validation component. These weights represent a conceptual design and cannot yet be positioned as the result of empirical optimization. The higher performance of the combined configuration compared with the individual models within the research environment is consistent with the direction of *ensemble* and *hybrid modelling* approaches (Gong et al., 2025; Zhou et al., 2025), but does not demonstrate universal superiority. SOP provides operational examination context after the model outputs become available, while SHAP provides post-prediction interpretation without altering the probabilities or *FinalScore*. Thus, the primary value of late-fusion in FDAS lies in integrating two different predictive perspectives within a modular structure to support decision-making, rather than claiming that one architecture is always superior to alternative approaches.

Explainable Artificial Intelligence (XAI) Analysis

In the context of FDAS, the practical value of SHAP lies in its ability to connect a *model alert* with information that can be reviewed by an analyst. At the individual case level, *local explanation* shows features that relatively increase or decrease the fraud probability, allowing analysts to compare the rationale behind a prediction with the transaction context before determining further action. As a conceptual illustration, when an alert shows relatively high contributions from transaction frequency, activity density, and temporal deviation, the analyst can examine whether the combination of these characteristics is consistent with the transaction history under review. This explanation serves as an initial examination aid, rather than evidence that a transaction is definitively fraudulent, and is consistent with the use of XAI to support user understanding of model outputs in *financial fraud detection* (Zhou et al., 2023).

At the aggregate level, *global explanation* through Mean Absolute SHAP helps identify features that are relatively dominant in the model's behavior and can be used as a basis for monitoring and reviewing prediction results. These values represent relative contributions rather than causal relationships. Within the FDAS investigation workflow, SHAP information forms a bridge between predictive outputs and operational examination: the model generates an alert, SHAP provides an interpretive basis, and the analyst then assesses its consistency with the transaction context and operational requirements before proceeding with the investigation. This mechanism can strengthen *transparency*, *traceability*, and *auditability* because the rationale supporting a prediction can be traced through feature contribution information, consistent with the literature's attention to explainability in AI systems that require openness and decision traceability (Alshudukhi et al., 2025; Nimmy et al., 2025).

SHAP contributions differ from predictive performance. SHAP does not improve *Accuracy*, *Precision*, *Recall*, *F1-score*, or *ROC-AUC*, nor does it alter the model probabilities, S_{SOP} , or *FinalScore*. Its added value lies in providing information that supports *human oversight* before an investigative decision is established. Because the study uses a synthetic dataset within a controlled experimental environment and FDAS remains an *operational prototype*, these benefits are positioned as a *decision-support demonstration*, rather than evidence of improved analyst effectiveness or validation in actual banking operations. Thus, the workflow **model alert** → **SHAP explanation** → **analyst interpretation** → **operational verification** → **human investigation** strengthens the decision-support function without replacing human judgment.

Research Comparison, Implications, and Limitations

The literature indicates that the development of *fraud detection* is increasingly moving toward data-driven modeling with different methodological focuses. (Chen et al., 2025) maps developments in *deep learning* for *financial fraud detection*, including challenges related to *imbalanced data*, *explainability*, and implementation, whereas (El Hachimi et al., 2026) provides direct evidence of the use of LSTM for *credit card fraud detection* under *imbalanced class* conditions. (Kim & Kim, 2025) and (Gong et al., 2025) are used as methodological comparators outside the fraud domain: (Kim & Kim, 2025) demonstrates the application of LightGBM to multivariate prediction while considering imbalance and *false-negative risk*, whereas (Gong et al., 2025) demonstrates the integration of temporal models and LightGBM through *ensemble learning*. Compared with these research strands, FDAS does not propose a new classification algorithm but brings together tabular-temporal modeling, imbalance handling, operational validation, and explainability within a single *decision support* framework (Chen et al., 2025; El Hachimi et al., 2026; Gong et al., 2025; Kim & Kim, 2025).

Table 12. Comparison of the Proposed FDAS with Recent Studies

Study	Main Approach	Imbalance / XAI	Decision-Support Orientation	Main Distinction
(Chen et al., 2025)	Systematic review of deep learning in financial fraud detection	Imbalance and explainability identified as major issues	Governance and implementation challenges	Research mapping, not a single prototype
(El Hachimi et al., 2026)	LSTM-based credit card fraud detection	Imbalanced-class handling	Primarily predictive	Directly related to fraud detection
(Kim & Kim, 2025)	Cost-sensitive LightGBM prediction	Imbalance and false-negative considerations	Risk-oriented prediction	Methodological comparator outside fraud
(Gong et al., 2025)	Temporal model + LightGBM ensemble	Ensemble integration	Prediction-oriented	Heterogeneous ensemble outside fraud
Proposed FDAS	LightGBM + LSTM + weighted late-fusion	TPH-SMOTE + KHOI-SMOTE + SHAP	SOP validation + Decision Support	Integrated decision-support prototype

The methodological implication of FDAS lies in the functional integration of several components with distinct roles: LightGBM and LSTM maintain independent tabular and temporal modeling pathways, *weighted decision-level late-fusion* combines the predictive outputs, TPH-SMOTE and KHOI-SMOTE form minority-class representations in the training data, SOP provides operational validation, and SHAP provides post-prediction explanations. This arrangement produces a more structured workflow for supporting analysts because predictive outputs do not stand alone but are accompanied by validation context and interpretive information. The contribution of the study therefore lies in workflow integration and process traceability, rather than in claiming that each component individually outperforms previous approaches.

Interpretation of the findings remains constrained by the characteristics of the data and the system's implementation level. The research dataset consists of synthetic data constructed from the characteristics of observed fraud and normal transaction patterns at several conventional banks, rather than actual customer transactions. This approach enables evaluation in a controlled environment without exposing sensitive information, but limits *external validity* and generalizability to more diverse operational conditions. The research unit is also limited to 50 accounts, comprising 25 fraud accounts and 25 normal accounts, so the behavioral heterogeneity that can be represented remains limited. FDAS was developed through Google AI Studio as a *prototype*, so its implementation results are more appropriately positioned as a *proof-of-concept* for *end-to-end* integration rather than evidence of readiness for *deployment* in an actual banking environment. The 40:40:20 *late-fusion* composition is also a conceptual design that has not undergone comprehensive optimization and *sensitivity analysis*.

Future research should employ larger and more heterogeneous datasets and conduct external validation using data that more realistically represent operational conditions. Further development should also include optimization and *sensitivity analysis* of *late-fusion* weights, a more formal *ablation study*, evaluation of SHAP explanation stability, and prototype testing with end users to assess FDAS consistency in more realistic investigative contexts.

CONCLUSION

This study develops the Fraud Detection and Analysis System (FDAS) as a Decision Support System through the integration of LightGBM for tabular representation and LSTM for temporal patterns. The two models are trained independently and combined through weighted decision-level late-fusion, with TPH-SMOTE and KHOI-SMOTE applied only to the *training set* to address data imbalance. SOP Validation is positioned as an *operational validation layer*, while SHAP is used as a *post-hoc explanation layer* to support interpretability, traceability, and *human oversight* without altering the model probabilities or *FinalScore*. In the final configuration, Accuracy of 97.4%, Precision of 96.8%, Recall of 96.2%, F1-score of 96.5%, and ROC-AUC of 0.984 were obtained based on 1,000 paired evaluation units, rather than 50 account-level decisions.

Comparison across configurations shows that the final configuration achieved the highest performance among the compared configurations, but this result is not interpreted as evidence of universal optimality or individual causal relationships attributable to each component. The final confusion matrix consists of 358 True Positive, 616 True Negative, 12 False Positive, and 14 False Negative, which is consistent with the evaluation of 1,000 units. The primary contribution of this study lies in integrating tabular-temporal modeling, *imbalance* handling, operational validation, and *post-hoc explainability* within a single prototype decision-support framework. The 40:40:20 *late-fusion* composition represents a conceptual design intended to balance the contributions of

LightGBM, LSTM, and SOP, rather than a configuration demonstrated to be optimal. FDAS supports analysts through recommendations and interpretive information but does not replace human decision-making or automatically block accounts.

These findings should be understood within the limits of the experimental environment used. The dataset consists of synthetic data constructed based on the characteristics of observed fraud and normal transaction patterns at several conventional banks, comprising 50 accounts, consisting of 25 fraud accounts and 25 normal accounts, over the period from January to December 2025. This configuration enables controlled methodological testing but limits behavioral diversity, external validity, and generalizability to broader operational conditions. FDAS was developed through Google AI Studio as a prototype, and therefore the results demonstrate a *proof-of-concept* of *end-to-end* integration rather than *deployment* readiness in an actual banking environment. Future research should expand the size and heterogeneity of the dataset, conduct external validation, examine optimization and *sensitivity analysis* of the *late-fusion* weights, strengthen the *ablation study*, evaluate the stability of SHAP explanations, and test the prototype under more realistic usage scenarios.

REFERENCES

- Almusallam, N., & Qayyum, J. (2025). A Hybrid Feature Selection and Clustering-Based Ensemble Learning Approach for Real-Time Fraud Detection in Financial Transactions. *Computers, Materials & Continua*, 85(2), 3653–3687. <https://doi.org/10.32604/cmc.2025.067220>
- Alshudukhi, K. S., Ali, S., Humayun, M., & Alruwaili, O. (2025). Next-Generation Lightweight Explainable AI for Cybersecurity: A Review on Transparency and Real-Time Threat Mitigation. *Computer Modeling in Engineering & Sciences*, 145(3), 3029–3085. <https://doi.org/10.32604/cmes.2025.073705>
- Bernardo, A., & Della Valle, E. (2022). An extensive study of C-SMOTE, a Continuous Synthetic Minority Oversampling Technique for Evolving Data Streams. *Expert Systems with Applications*, 196, 116630. <https://doi.org/10.1016/j.eswa.2022.116630>
- Chen, Y., Zhao, C., Xu, Y., Nie, C., & Zhang, Y. (2025). Deep Learning in Financial Fraud Detection: Innovations, Challenges, and Applications. *Data Science and Management*, S2666764925000372. <https://doi.org/10.1016/j.dsm.2025.08.002>
- Du, Z., Yan, S., Wang, J., Guo, X., & Liu, S. (2026). TPH-SMOTE: A tri-process heuristic oversampling approach integrating SMOTE and whale optimization for imbalanced binary classification. *Expert Systems with Applications*, 306, 130952. <https://doi.org/10.1016/j.eswa.2025.130952>
- El Hachimi, K., Orhanou, G., & Ballihi, L. (2026). Credit card fraud detection model based on optimized long short-term memory: Imbalanced class solution and overfitting reduction. *Expert Systems with Applications*, 305, 130871. <https://doi.org/10.1016/j.eswa.2025.130871>
- Gong, J., Qu, Z., Zhu, Z., Xu, H., & Yang, Q. (2025). Ensemble models of TCN-LSTM-LightGBM based on ensemble learning methods for short-term electrical load forecasting. *Energy*, 318, 134757. <https://doi.org/10.1016/j.energy.2025.134757>
- Kim, S.-W., & Kim, Y. S. (2025). Two-Stage LightGBM Framework for Cost-Sensitive Prediction of Impending Failures of Component X in Scania Trucks. *Computers, Materials & Continua*, 0(0), 1–10. <https://doi.org/10.32604/cmc.2025.073492>
- Liu, X., Xie, S., Yang, J., Wu, S., Du, M., Ji, Y., Zheng, H., An, H., Zhou, W., Lv, H., Ji, G., Yang, Q., Xu, L., Chen, Y., & Yang, Y. (2026). KHOI-SMOTE: An efficient oversampling technique based on k-means clustering and h-outlyingness index for imbalanced medical data. *Neurocomputing*, 671, 132643. <https://doi.org/10.1016/j.neucom.2026.132643>
- Lukács, A., & Váradi, S. (2023). GDPR-compliant AI-based automated decision-making in the world of work. *Computer Law & Security Review*, 50, 105848. <https://doi.org/10.1016/j.clsr.2023.105848>
- Nimmy, S. F., Hussain, O. K., Chakraborty, R. K., & Saha, S. (2025). Explainable Artificial Intelligence (XAI) in glaucoma assessment: Advancing the frontiers of machine learning algorithms. *Knowledge-Based Systems*, 316, 113333. <https://doi.org/10.1016/j.knsys.2025.113333>
- Uddin, N. (2025). Role of AI in Preventing Financial Crime: A Comprehensive Analytical Review. *Journal of Economic Criminology*, 100200. <https://doi.org/10.1016/j.jeconc.2025.100200>
- Widodo, A. O., Setiawan, B., & Indraswari, R. (2024). Machine Learning-Based Intrusion Detection on Multi-Class Imbalanced Dataset Using SMOTE. *Procedia Computer Science*, 234, 578–583. <https://doi.org/10.1016/j.procs.2024.03.042>
- Yousefimehr, B., & Ghatee, M. (2025). A distribution-preserving method for resampling combined with LightGBM-LSTM for sequence-wise fraud detection in credit card transactions. *Expert Systems with Applications*, 262, 125661. <https://doi.org/10.1016/j.eswa.2024.125661>
- Zhou, Y., Li, H., Xiao, Z., & Qiu, J. (2023). A user-centered explainable artificial intelligence approach for financial fraud detection. *Finance Research Letters*, 58, 104309. <https://doi.org/10.1016/j.frl.2023.104309>

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Zhou, Y., Wang, S., Li, Z., Feng, R., & Fernandez, C. (2025). Battery pack capacity estimation based on improved cooperative co-evolutionary strategy and LightGBM hybrid models using indirect health features. *Journal of Energy Storage*, 131, 115914. <https://doi.org/10.1016/j.est.2025.115914>